Is the world becoming a happier place?

Aaron Lin, Siddharth Manu, Stephanie Andrews

DataSci 200 Fall 2023

Primary Question

What factors contribute to global happiness?

Sub-questions

Q1   Is the world becoming a happier place?

Q2   Are countries with greater wealth inequality less happy?

Q3   Are wealthier countries happier?

Q4   Are countries with higher climate risk unhappier?

Q5   Is there a correlation between clean air and happiness?

Project Goals
- Visualize global happiness levels, by country and by region
- Analyze factors that may contribute to changes in national happiness
- Provide actionable insights into the most important factors to improve quality of life

Visualize global happiness levels, by country and by region and explore some factors that may contribute to changes in national happiness. Our sub-questions reflect this and are as follows:

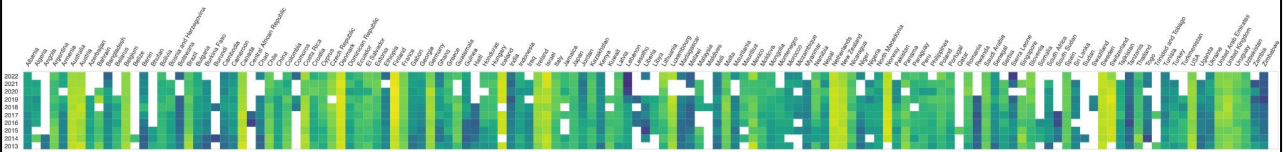## Datasets

Primary:
World Happiness Report Data

Supplemental:
- World Bank Population Data
- Environmental Performance Index
- Climate Risk Index
- World Inequality Database
- World Map GeoJson

WHR 2023

Context

- Primary source for WHR: Gallup World Poll
  - Approx. 1,000 responses annually from individuals in each country
- Life Evaluations via Cantril Life Ladder (LL) Scores from Gallup Poll
  - Respondents rate their life on a scale from 0 (worst) to 10 (best)
  - Historical max: 8.02; Historical min: 1.28
- WHR (Sustainable Development Solutions Network) uses both LL Scores and data on positive emotions and negative emotions
- WHR then generates national rankings
- Our analysis:
  - Examines data without explicit ranking
  - Scoped to LL Scores only
  - Explores other potential contributing factors

**Average Life Evaluation Scores**

Our primary dataset is of global happiness scores collected via the annual Gallup World Poll. For this poll Gallup collects one thousand responses annually from individuals in each country, life evaluations among them as a measure of subjective well-being. On these surveys, respondents rate their life on a scale of 0 (being the worst) to 10 (being the best) The historical max score is 8.02, and the historical minimum score is 1.28. The World Happiness Report uses these scores to generate national happiness rankings. The World Happiness Report also uses data collected about positive emotions and negative emotions.

In our analysis, we examine the raw life ladder scores to explore national happiness levels. Unlike the World Happiness Report, however, we do so without explicitly ranking countries, have scoped our analysis specifically to life ladder scores, and also try to explore potential contributing factors like climate and wealth inequality. But first, before going into that, we wanted to know whether the world is actually becoming a happier place.

We created a few heatmaps to visualize and explore life ladder scores over the past 10 years. This first one at the bottom displays Life Ladder Scores for every country during this period, sorted alphabetically.
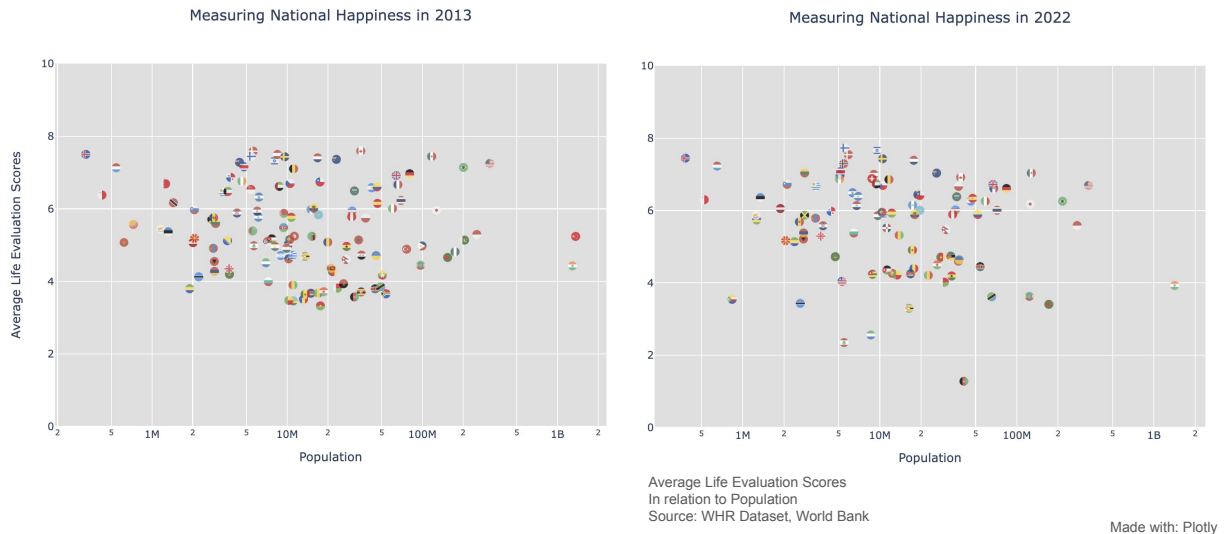
Q1 Is the world becoming a happier place?

Average Life Evaluation Scores
Grouped by Region
Source: WHR Dataset

Made with: Bokeh
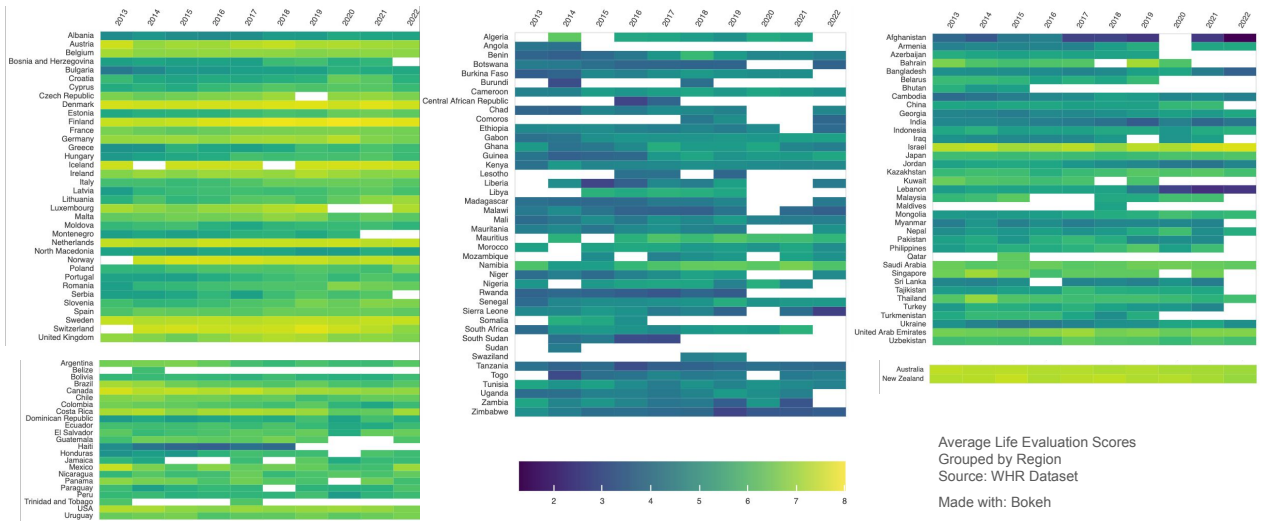
We grouped this data by region to make it easier to explore. Unsurprisingly, countries in Europe and the Americas reported higher overall happiness scores across most years, and countries in Africa and parts of Asia tended to report lower scores. The scores appeared to be generally trending downward over the years, which seemed to indicate a decrease in overall national happiness.

We then created a few snapshots comparing two years.

One decade of change: a snapshot

Measuring National Happiness in 2013

Measuring National Happiness in 2022

Average Life Evaluation Scores
In relation to Population
Source: WHR Dataset, World Bank

Made with: Plotly

First is a scatter plot representing a decade of change, with scores on the y-axis and population on the x-axis. The scatter plot on the left visualizes data from 2013; on the right is data from the most recent year. There's a somewhat central cluster of scores in 2013. By 2022, there is a greater spread, and a number of countries have reported much lower happiness scores.
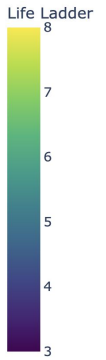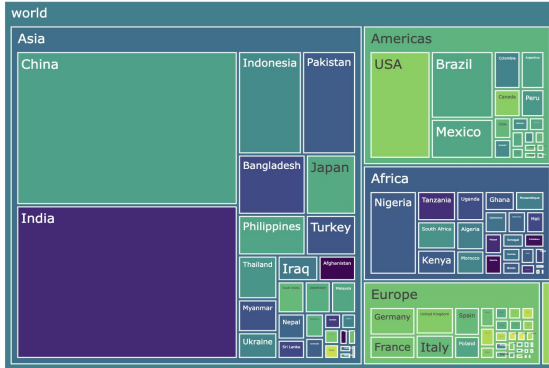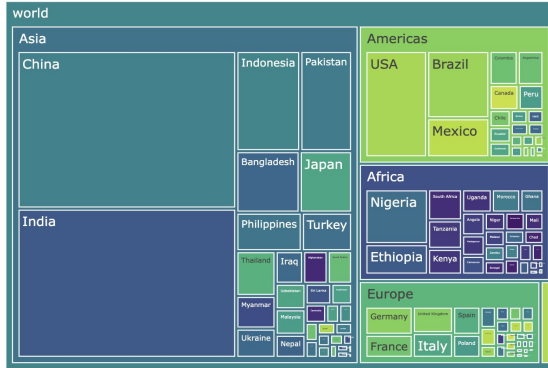
We also noticed that we could not find China in 2022. It seems China did not participate in the most recent survey.

So we took snapshots of life ladder scores in 2013 against the last year we had data from China, which was 2021. These treemaps illustrate life ladder scores in relation to population. Each block is a country, with size proportional to population and color scaled to the life ladder score. Again, brighter scores indicate higher happiness levels.

This data also indicates a decrease in global happiness in the past decade.

==Another snapshot: pre-pandemic data==

Average Life Evaluation Scores
Grouped by Region
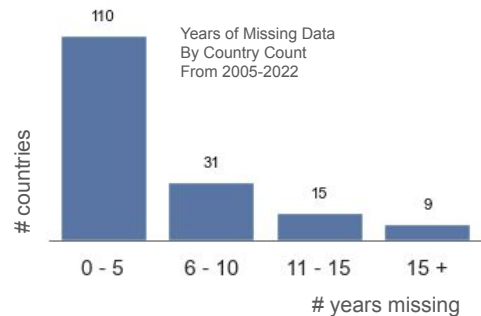Source: WHR Dataset
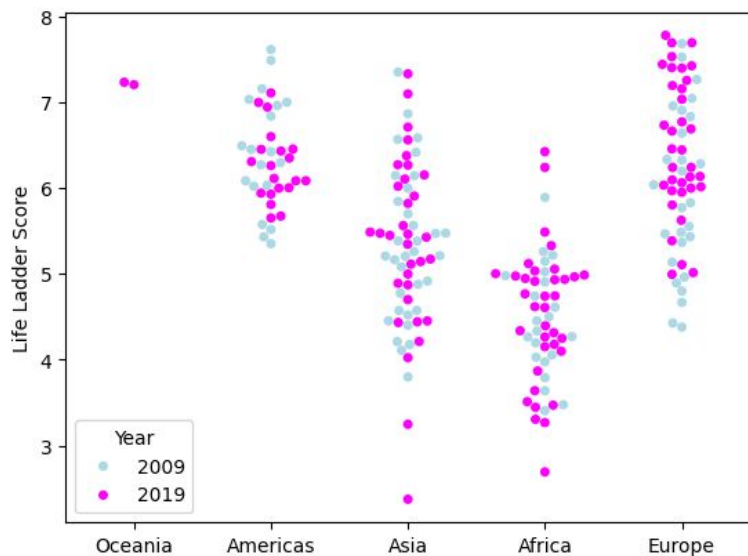
Made with: Seaborn

But given that one year is during the pandemic, we wanted to also look at how national happiness levels may have changed due to factors unrelated to the pandemic. So we created a swarm plot comparing scores from the year right before the pandemic, 2019, against scores from a decade prior. These are grouped by region to make it easier to see trends.

In general, it does appear that the world is becoming a less happy place.

In our subsequent analyses, we try to look at some factors that may have contributed to fluctuations in global happiness, to identify additional areas where humanity can focus its efforts.

Pre-Processing, Data Issues, and Assumptions

- WHR Data
  - Not all countries participated
  - Not all years available for each participating country

- Climate risk
  - Country names don't match
  - Not all years available for each country
- Air quality
  - Aggregate datasets for different factors based on their ratios

Years of Missing Data
By Country Count
From 2005-2022

110

31

15

9

\# countries

0 - 5   6 - 10   11 - 15   15 +

\# years missing

Our team had to handle many issues working with our datasets. In the interest of time, we'll cover a few of the bigger ones we encountered.

First, with regards to data from the World Happiness Report, not all countries participated in the annual surveys. And data was not available for some years among the countries that did participate.

This bar chart shows the years of missing data by country count. The x-axis corresponds to number of years missing per country, and the y-axis corresponds to the count of countries that were missing those number of years.

With the climate risk dataset, which Sid will discuss in further detail later, country names did not match. And like the happiness report data, data was not available for some countries in some years.

For the air quality dataset, our team had to aggregate datasets for different factors based on their ratios, which was a fairly involved process.

X 778 files

Adding on to what Stephanie mentioned, the WID data proved to be challenging. Due to issues with exporting filtered data from the website, it was necessary to extract and process the data directly. There were 778 files consisting of raw data files and metadata files for unique countries and regions. File paths were obtained and filtered using the python os library and regex. Pandas was used to load the csv files into data frames and join the respective metadata and raw data dataframes for each country. These were then concatenated to obtain one dataframe with columns of the desired variable names, variable values, percentiles, countries, and years.

Q2 Happiness and Wealth inequality: are they related?

The second question our analysis sought to answer is whether or not Happiness values as reported by the WHR exhibits any relationships with wealth inequality. To answer this question, the World Inequality Database was used to extract the percentage share of private wealth owned by the top 10% and bottom 50% of the population. From these values, a ratio was calculated with higher values indicating greater wealth inequality. The median values for each country's time series data were plotted in a bar plot by country in descending order of Log Wealth Inequality Ratios with bar colors representing each country's corresponding happiness index values. Darker shades of blue indicate happier countries. This graph visualizes a relatively even distribution of happiness values across the spectrum, which suggests that there may not be a strong correlation between happiness and wealth inequality from the selected metrics. But let's try visualizing this data in another way.

Happiness is not an indicator of Wealth Inequality Ratio

Life Ladder vs. Log Wealth Inequality Ratio
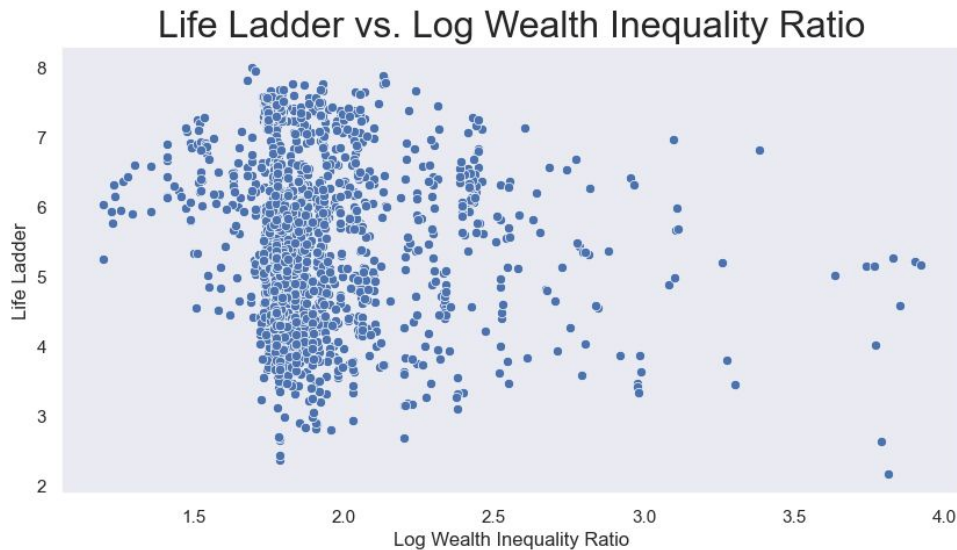
The next visualization of this data is a scatter plot with the Happiness Values represented by the World Happiness Report's Life Ladder index on the Y-axis and the Log Wealth Inequality Ratio on the X-axis. This graph was prepared by plotting every country's data from 2005 to 2022 due to data availability. One notable outlier was Lebanon in the year 2011, which had a ratio of over 90,000 represented by an x-value of around 5 in this plot. This value was excluded due to the context of the year in the country's history; 2011 was a year of great turmoil in the country due to protests and the creation of a new government. In examining this scatter plot, we can see there is no discernable relationship between the selected variables though further analysis is necessary to create confidence in this claim.

Q3 Does Country Wealth Indicate Happiness?

Happiness Ladder Score vs. Log GDP Per Capita
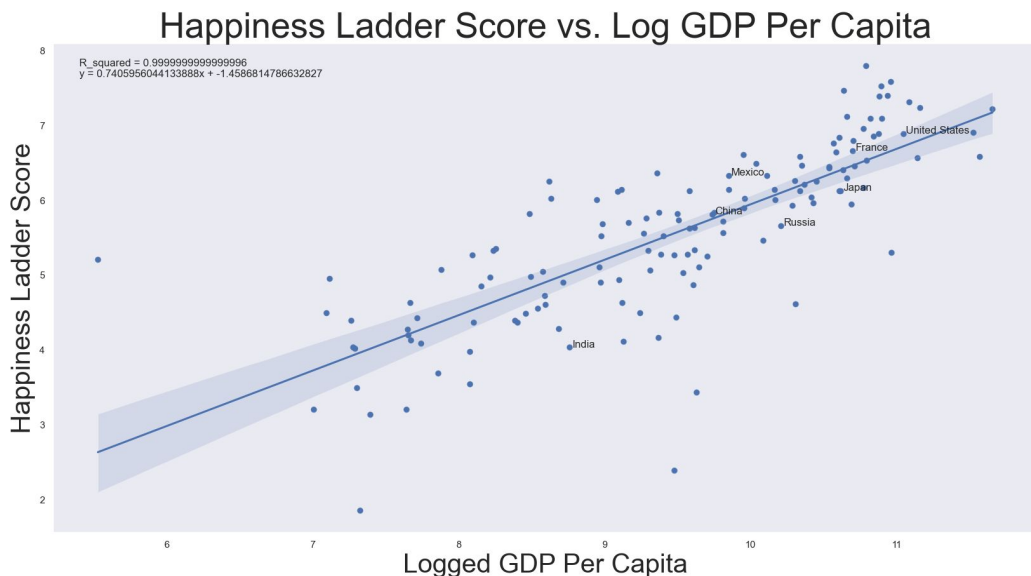
The third question this report sought to investigate is whether or not a country's wealth was an indicator of happiness. For this analysis, a country's wealth was represented by its Logged GDP Per Capita visualized on the X-axis with the Happiness Ladder Score on the Y-axis. The plot indicates a strong positive correlation between the two variables, which was confirmed by a linear regression that indicated that they had an R-squared value of over .99 as shown in the top left of the graph. Despite the strength of the correlation, more rigorous statistical analysis is necessary to determine if these variables have a causal relationship.

For the next part, I'd like to pass it on to my colleague, Sid.

Climate change is a global risk

Climate Risk by Country
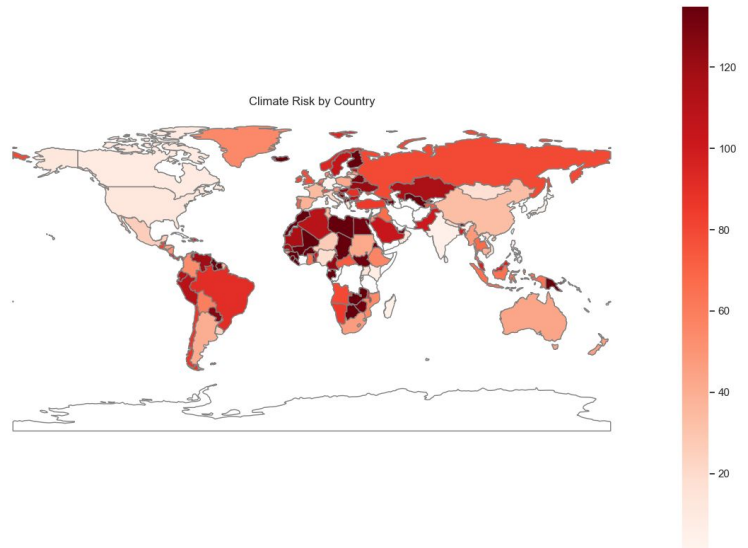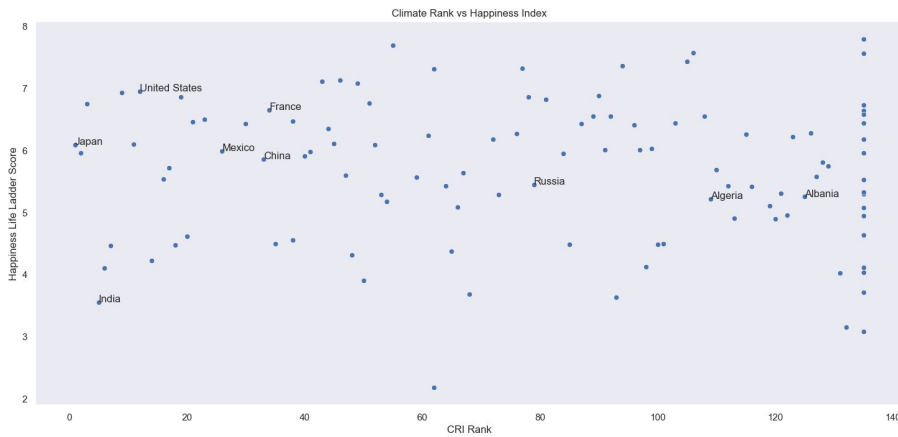
Thank you Aaron!

The next dataset we looked into was the Climate Risk Index. This index is a tool that assesses and ranks countries based on their vulnerability to climate change impacts, considering factors such as extreme weather events. It helps identify nations at higher risk and informs strategies for climate resilience and adaptation.

The world map above shows Climate Risk by country for the year 2018. The darker shades indicate higher climate risk.
One thing to note here is that we do not have all the countries participating and hence we don't have data for them. They are shown as white.

Well, you can opt-out of surveys for climate change but not climate change itself. :)

Q4 Climate Risk vs Happiness Index

Climate Rank vs Happiness Index

Average Life Evaluation Scores
In relation to Climate Risk
Source: WHR Dataset, CRI

No
correlation???

Or are we oblivious to the future?

In order to delve into the question of climate risk and happiness index of a nation, we created a scatter plot with the CRI rank on the X-axis and the Happiness score on the Y-axis. While we expected countries at higher risk to be unhappier, we couldn't find a strong correlation.
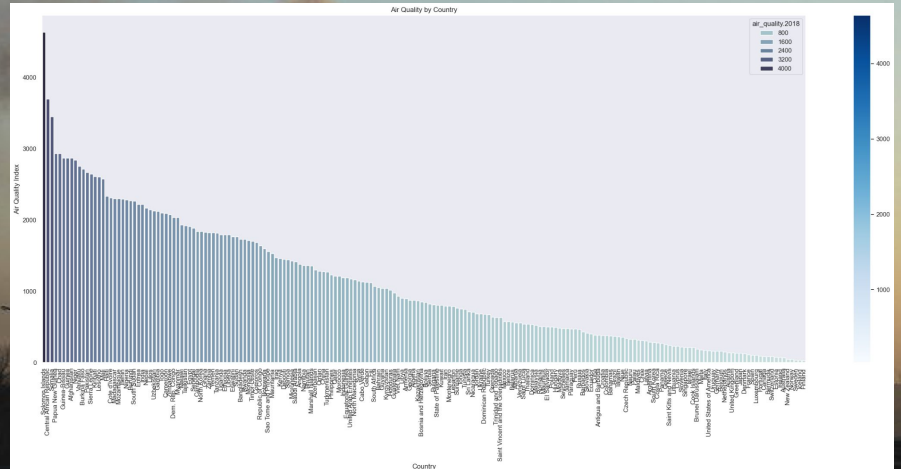
People even in high risk countries seem to be relatively happy. Which should make us think - are we so oblivious to the future? :)

Air Quality

PM$_{2.5}$ Exposure	PMD	47%
Household Solid Fuels	HAD	38%
Ozone Exposure	OZD	5%
NOx Exposure	NOE	5%
SO2 Exposure	SOE	2%
CO Exposure	COE	2%
VOC Exposure	VOE	2%

Weights in Air Quality calculation

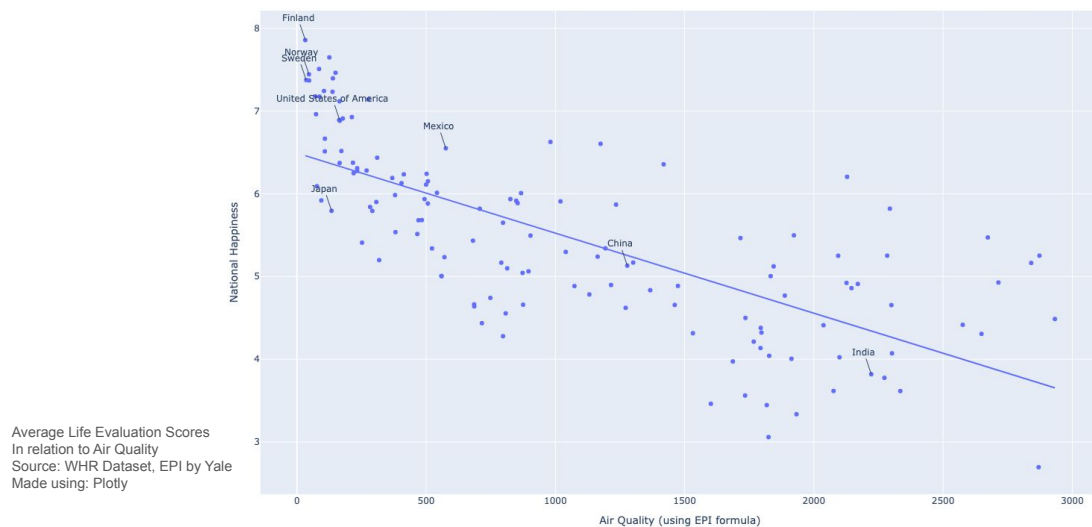Air quality varies HUGELY among countries.

Air Quality (EPI formula)
Source: EPI by Yale
Made using Seaborn

The final question we pursued about the relationship between clean air quality and happiness index.

In order to calculate the air quality, we merged datasets for each of the individual factors like SO2, CO exposure etc based on country and then calculated the air quality index using the weight ratio for each factor used in the PFI report. We also filtered out countries for which the air quality score was calculated to be negative, which is an invalid value.

The above bar chart has countries on the X-axis and air quality on the Y-axis. A lower air quality score here indicates better air quality. As we can see in the plot, air quality varies greatly among countries.

Q5 Air Quality vs Happiness?

Chart axes: National Happiness (Y-axis) vs Air Quality (using EPI formula) (X-axis)

Labeled countries: Finland, Norway, Sweden, United States of America, Mexico, Japan, China, India

Average Life Evaluation Scores
In relation to Air Quality
Source: WHR Dataset, EPI by Yale
Made using: Plotly

To study the relationship of air quality with happiness, let's look at this scatter plot with air quality on the X-axis and national happiness on the Y-axis. The regression line drawn here uses the OLS technique.

We can clearly see a strong correlation between clean air and happiness index. Countries like Finland and Norway are at the better end on both indices!

Implications and Actionable Insights

- Analysis suggests that correlations MAY exist for:
    - GDP
    - Air Quality
- Further Analysis
    - Other environmental factors
    - Conflict, terror
    - Technical innovations
    - Social factors

In conclusion, our analysis reveals strong correlations between a country's happiness index and factors such as GDP and air quality.

To further our understanding of world happiness, the next steps would be conducting further analysis on additional environmental factors like clean water, examining correlations with conflict and terror, and exploring the impact of technical innovations on the state of the world. Additionally, delving into social factors could uncover nuanced influences on national happiness.

These insights can guide policymakers and stakeholders in formulating targeted strategies to enhance well-being, focusing on areas like environmental policies, social programs, and innovation initiatives. This holistic approach ensures a comprehensive understanding and effective interventions to foster greater happiness in countries, ultimately, for a happier world!
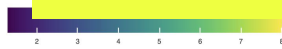
References

- [World Happiness Report](#)
- [World Happiness Report Data](#)
- [Gallup](#): "Understanding How Gallup Uses the Cantril Scale"
- [World Bank Population Data](#)
- [Environmental Performance Index](#)
- [Climate Risk Index](#)
- [World Inequality Database](#)
- [World Map GeoJson](#)
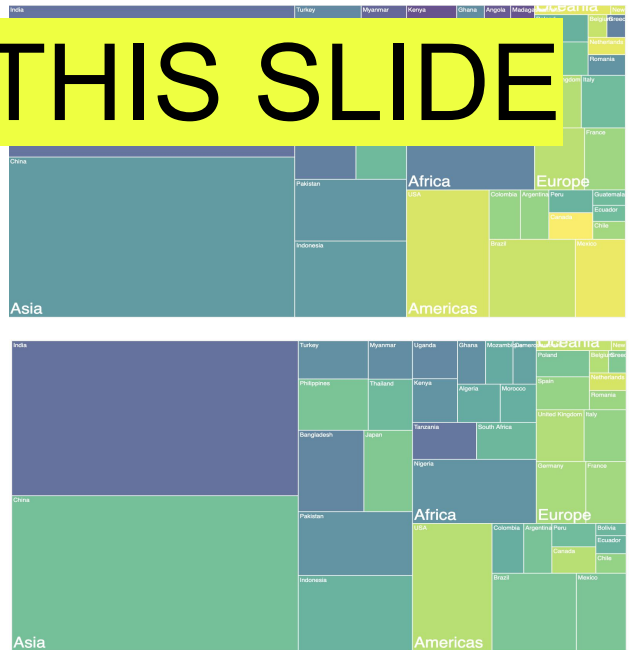
Thank you so much for listening!

We would now like to open the floor for questions.

Made with: Bokeh

Future viz?
- ● Non-tree map
- ● Global averages, weighted, per year (5, 6)
- ● Maybe scale swarm dots to pop
- ● UN Regions
  - ○ https://population.un.org/wpp/DefinitionOfRegions
  - ○ https://www.un.org/dgacm/en/content/regional-groups
  - ○ https://unstats.un.org/sdgs/indicators/regional-groups/

Resources
https://seaborn.pydata.org/generated/seaborn.FacetGrid.html
https://seaborn.pydata.org/tutorial/axis_grids.html
https://blog.datawrapper.de/cartograms/